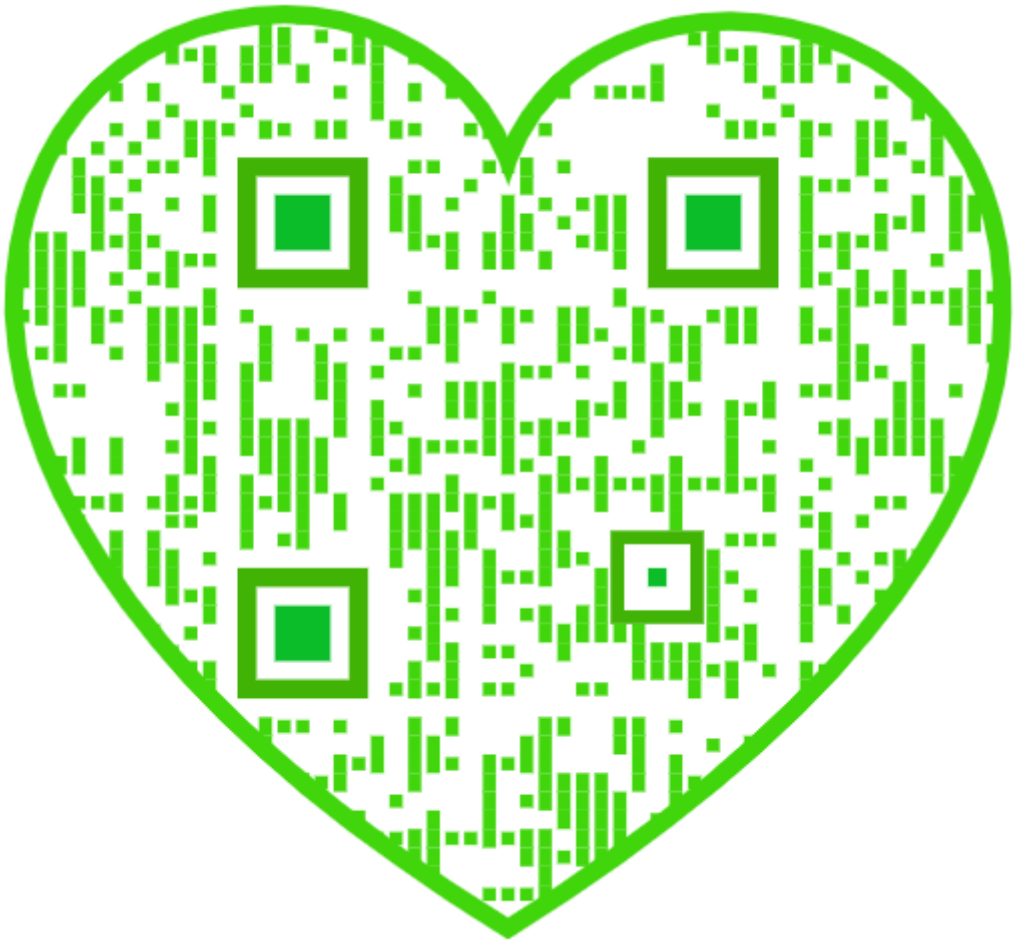
Master in Artificial Intelligence

Algorithm Selection & Development V



Purpose

The purpose of the section is to help you learn how to research, select, and develop appropriate algorithms to become a Successful Artificial Intelligence (AI) Engineer

At the end of this lecture, you will learn the following

- **How does Davies-Bouldin Index compute the average similarity between each cluster and its most similar cluster**

How to determine type of output and evaluation metrices?

Understand the problem

- Domain
- Objectives
- Constraints

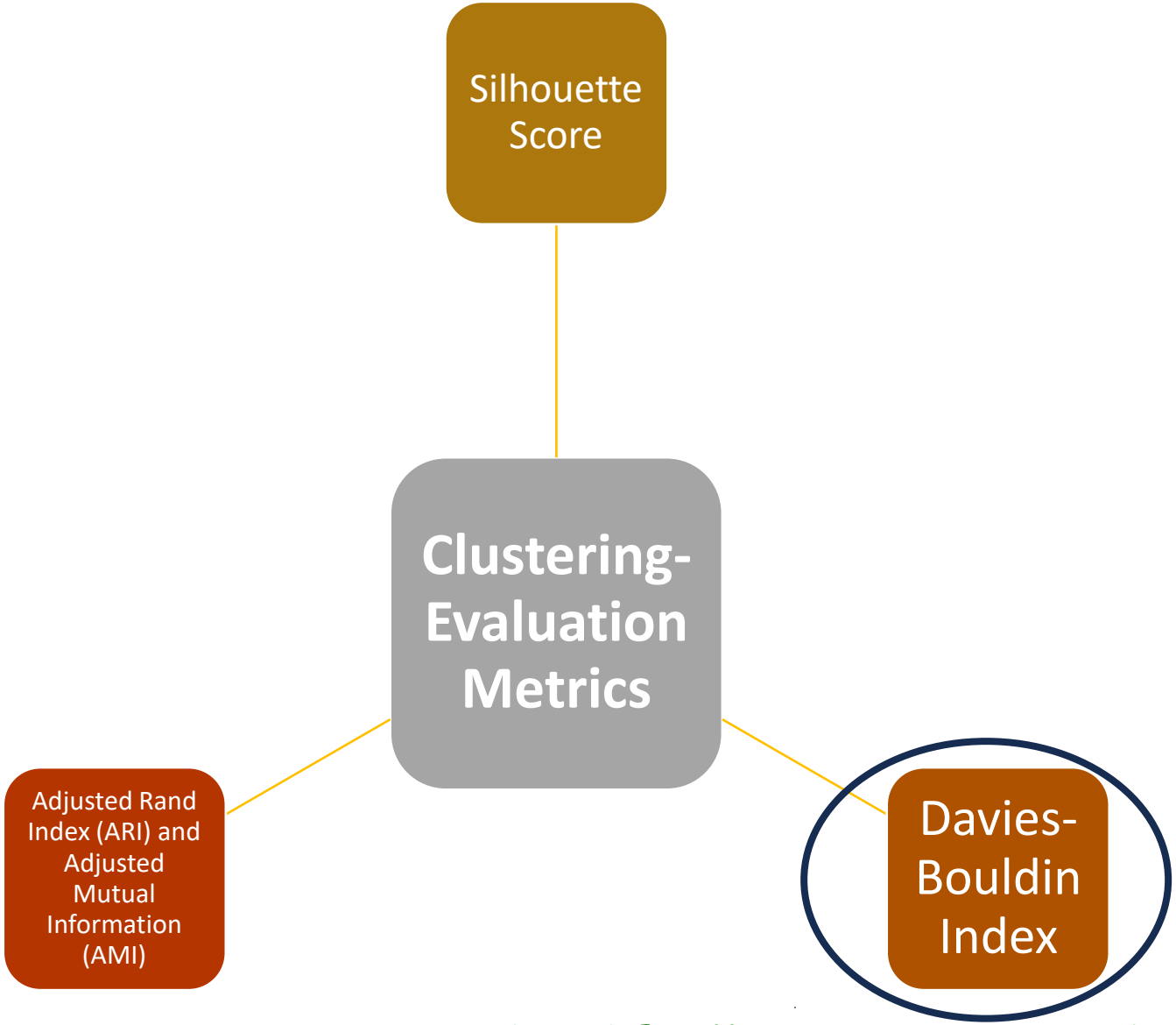
Define the problem as a

- Supervised
- Unsupervised
- Reinforcement learning task.

Determine

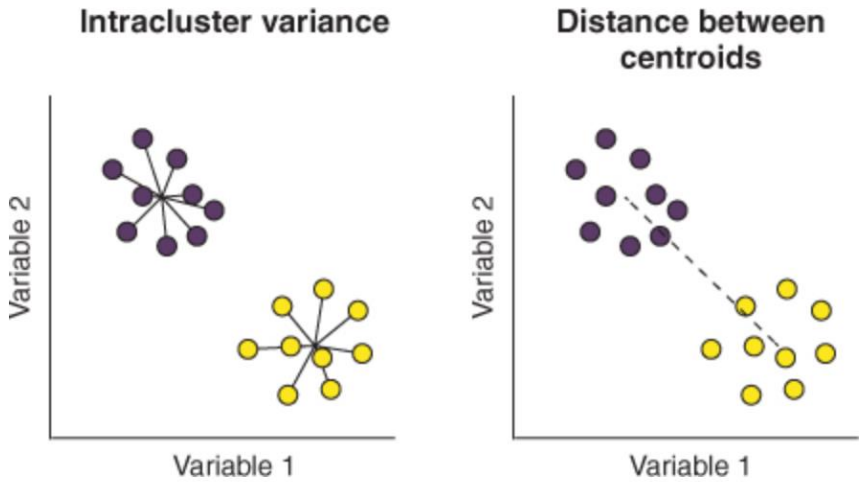
- type of output (e.g., classification, regression, clustering)
- Evaluation metrics

How does Davies-Bouldin Index compute the average similarity between each cluster and its most similar cluster



Calculating the Davies-Bouldin Index

- For each cluster *i*:
  - Compute the centroid *Ci*, which represents the center of the cluster.
  - For each other cluster *j* (where j≠*i*):
    - Calculate the average distance between each point in cluster *i* and the centroid *Ci*. This is denoted as avg*i*.
    - Calculate the average distance between each point in cluster *j* and the centroid *Cj*. This is denoted as avg*j*.
    - Compute the distance between the centroids *Ci* and *Cj*. This is denoted as *dij*.
  - Calculate the similarity between cluster *i* and its most similar cluster *j* as *Rij*=(avg*i*+avg*j*)/*dij*
- The Davies-Bouldin Index for the clustering is the average of the similarity values *Rij* across all clusters



Interpreting the Davies-Bouldin Index

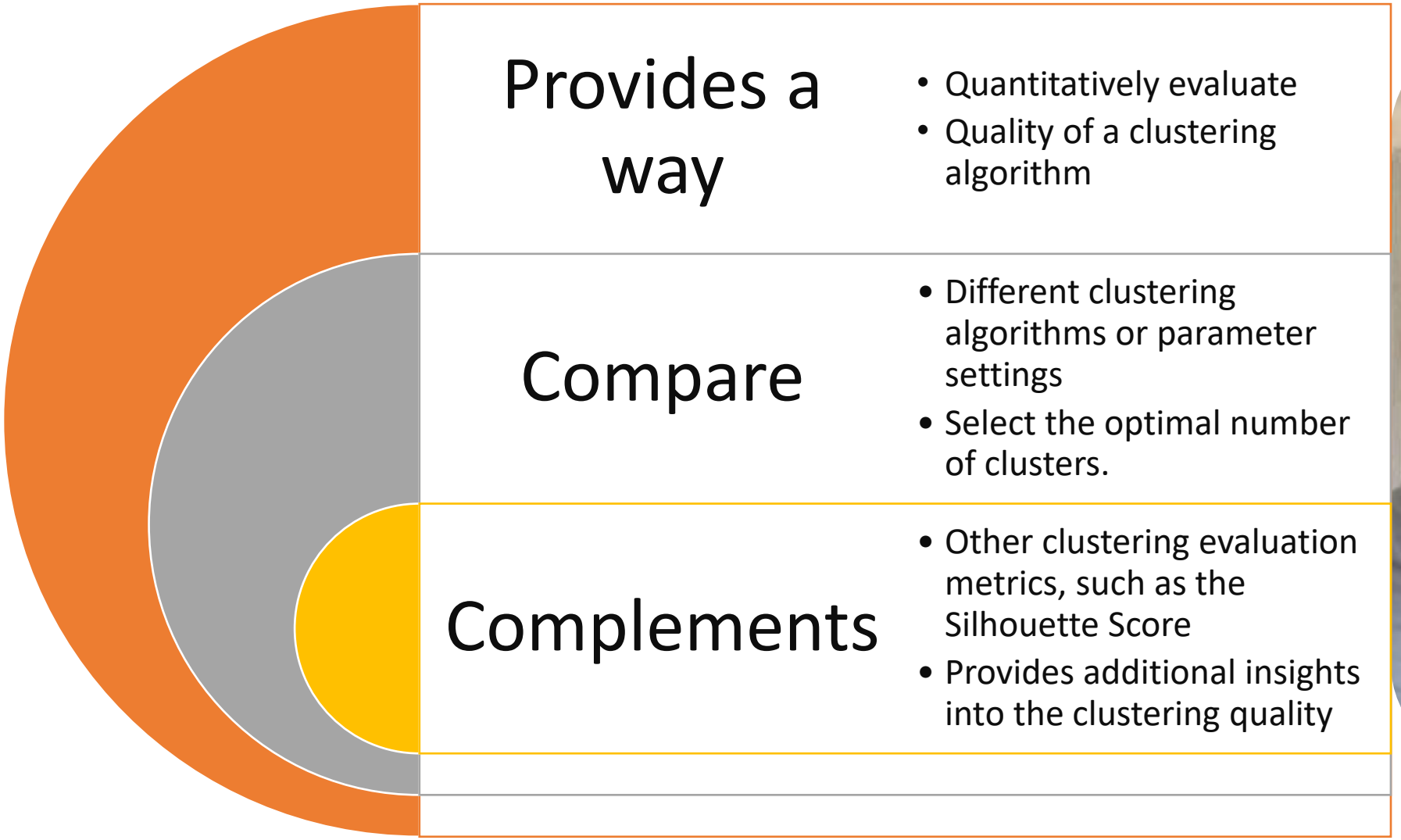
Low Davies-Bouldin Index

- Clusters are well-separated and distinct from each other
- Each cluster is more similar to its own cluster than to others

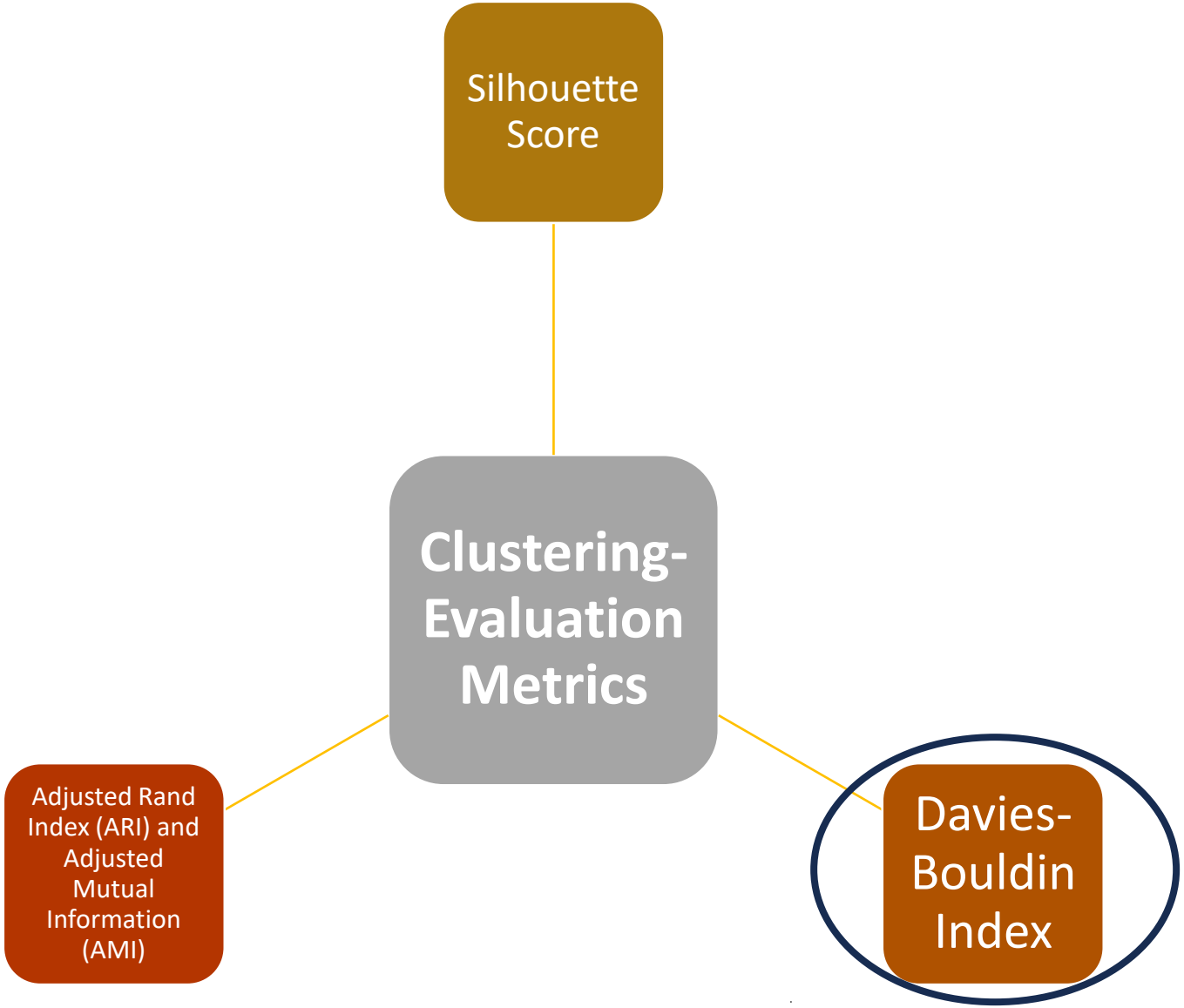
High Davies-Bouldin Index

- Clusters are less well-separated and may have significant overlap
- The boundaries between clusters are less clear

Using the Davies-Bouldin Index for Cluster Evaluation

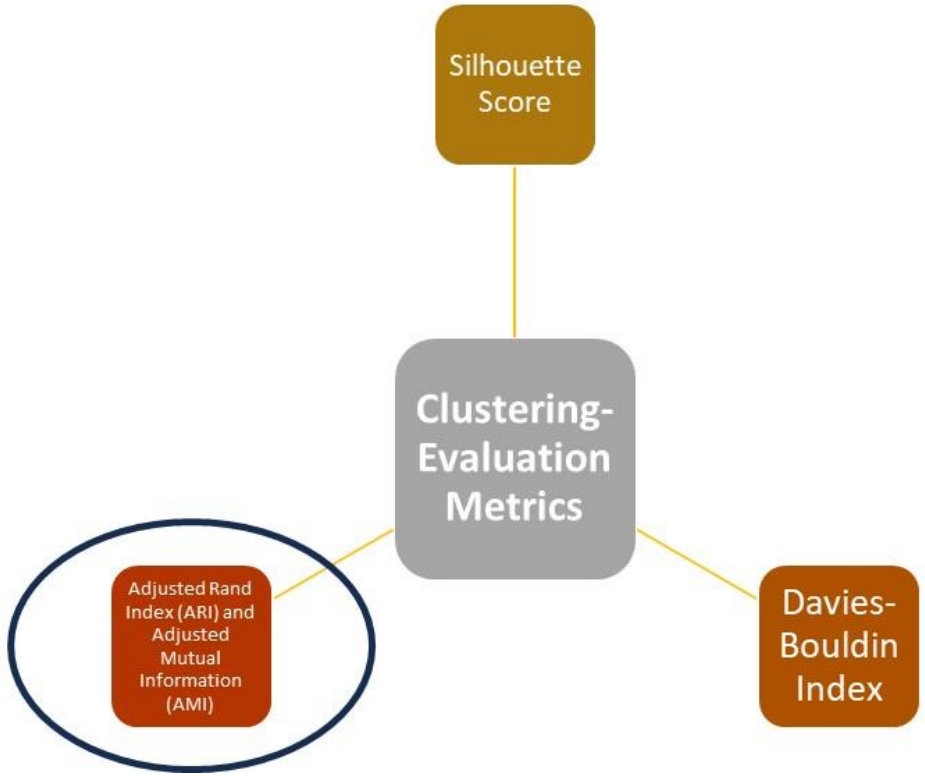


How does Davies-Bouldin Index compute the average similarity between each cluster and its most similar cluster



What is next?

Adjusted Rand Index (ARI) and Adjusted Mutual Information (AMI) Measures the agreement between true labels and cluster assignments



Master in Artificial Intelligence

Algorithm Selection & Development V